

UNIVERSITY OF WOLLONGONG

DOCTORAL THESIS

---

**Thesis Title**

---

*Author:*  
Jesse GREENSLADE

*Supervisor:*  
Dr. Jenny FISHER

*A thesis submitted in fulfillment of the requirements  
for the degree of Doctor of Philosophy  
in the*

Centre of Atmospheric Chemistry  
Chemistry Department

April 10, 2018



## Declaration of Authorship

I, Jesse GREENSLADE, declare that this thesis titled, “Thesis Title” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

---

Date:

---



*"Thanks to my solid academic training, today I can write hundreds of words on virtually any topic without possessing a shred of information, which is how I got a good job in journalism."*

Dave Barry



# Contents

<b>Declaration of Authorship</b>	<b>iii</b>
<b>Abstract</b>	<b>vii</b>
<b>Acknowledgements</b>	<b>ix</b>
<b>1 Introduction and Literature Review</b>	<b>1</b>
1.1 The atmosphere . . . . .	1
1.1.1 Structure . . . . .	2
1.1.2 Composition . . . . .	3
Hydroxyl radicals . . . . .	3
1.2 Ozone . . . . .	3
NO <sub>x</sub> . . . . .	4
1.2.1 Stratosphere to troposphere transport . . . . .	6
1.2.2 Chemical production . . . . .	7
1.3 VOCs . . . . .	7
1.3.1 Emissions . . . . .	9
1.3.2 Isoprene . . . . .	10
1.3.3 The isoprene cascade . . . . .	11
Oxidation . . . . .	11
Low NO <sub>x</sub> scenario . . . . .	12
Ozonolysis . . . . .	13
Oxidation by NO <sub>3</sub> . . . . .	13
1.3.4 Radiative Forcing . . . . .	14
1.4 HCHO . . . . .	14
1.4.1 Sources and sinks . . . . .	14
1.4.2 Measurement techniques . . . . .	17
Satellite measurements . . . . .	18
1.5 Atmospheric Chemistry Modelling . . . . .	19
1.5.1 Box models . . . . .	21
1.5.2 Emissions . . . . .	21
1.5.3 Uncertainties? . . . . .	22
Emissions Inventories . . . . .	22
Resolution . . . . .	23
Chemistry mechanisms . . . . .	23
Clouds . . . . .	24
Soil Moisture . . . . .	24
1.6 Australia and the southern hemisphere . . . . .	24
1.6.1 Ozone . . . . .	25

1.6.2	VOCs . . . . .	26
1.6.3	Measurements . . . . .	27
1.7	Aims . . . . .	28
<b>2</b>	<b>Data and Modelling</b>	<b>29</b>
2.1	Introduction . . . . .	29
2.2	List of runs and outputs used in my work TODO: good place for this? . . . . .	29
2.2.1	GEOS-Chem . . . . .	29
2.2.2	CAABA/MECCA . . . . .	30
2.2.3	Reading Data . . . . .	30
	CAABA/MECCA outputs . . . . .	30
	GEOS-Chem Satellite output . . . . .	30
	HEMCO diagnostics . . . . .	30
2.3	GEOS-Chem . . . . .	30
2.3.1	GEOS-Chem isoprene modelling . . . . .	31
	Outline . . . . .	31
	Emissions from MEGAN . . . . .	33
2.3.2	Chemical Mechanisms . . . . .	35
2.3.3	Running GEOS-Chem (before isop?) . . . . .	35
	Installation and requirements . . . . .	35
	Outputs . . . . .	35
	Tropospheric chemistry run . . . . .	36
	UCX run . . . . .	36
2.3.4	Run comparisons . . . . .	36
2.4	Measurement Techniques . . . . .	36
2.4.1	DOAS . . . . .	40
2.4.2	Satellites . . . . .	40
	OMI . . . . .	42
	AMF . . . . .	44
	Uncertainties . . . . .	44
2.4.3	OMI Satellite measurements . . . . .	46
	OMI HCHO . . . . .	46
	OMI NO <sub>2</sub> . . . . .	46
2.5	Campaigns . . . . .	46
2.5.1	Marine and Urban MBA ? (MUMBA) . . . . .	47
2.5.2	Sydney Particle Studies (SPS1, SPS2) . . . . .	47
2.6	Analysing output . . . . .	47
2.6.1	Circadian emissions cycle . . . . .	47
2.6.2	HCHO: Simulated vs Measured . . . . .	47
2.6.3	Accounting for Fires . . . . .	48
2.6.4	Accounting for NO <sub>x</sub> . . . . .	48
2.6.5	HCHO Comparisons . . . . .	53
2.7	Data Access . . . . .	53



<b>3</b>	<b>Biogenic Isoprene emissions in Australia</b>	<b>63</b>
3.1	Introduction	63
3.1.1	Top-down emissions estimates	64
	Linear	64
	Bayesian	64
3.1.2	Aims	65
3.2	Methods	65
3.2.1	Outline	65
3.2.2		66
3.2.3	Satellite inversion	66
3.2.4	Calculation of Emissions	69
3.2.5	Emissions drivers	70
3.2.6	HCHO Products and yield	70
3.2.7	Accounting for smearing	73
3.3	Results	76
3.3.1	Emissions comparisons	76
3.3.2	Emissions affect on GEOS-Chem	77
3.4	Uncertainty	77
3.4.1	Model Uncertainty	77
3.4.2	Satellite Uncertainty	77
<b>3</b>	<b>Stratospheric ozone intrusions</b>	<b>63</b>
3.1	Introduction	63
3.2	Data and Methods	65
3.2.1	Ozonesonde record in the Southern Ocean	65
3.2.2	Model description	68
3.2.3	Characterisation of STT events and associated fluxes	69
3.2.4	Biomass burning influence	71
3.2.5	Classifying synoptic conditions during STT events	72
3.3	STT event climatologies	72
3.4	Simulated ozone columns	75
3.5	Stratosphere-to-troposphere ozone flux from STT events	80
3.5.1	Method	80
3.5.2	Results	82
3.5.3	Comparison to literature	82
3.6	Sensitivities and limitations	87
3.6.1	Event detection	87
3.6.2	Flux calculations	88
3.7	Conclusions	89
3.8	Contributions and Acknowledgements	90
<b>5</b>	<b>TODO: move to biogenic isop chapter: Formaldehyde product over Australia</b>	<b>111</b>
5.1	Australian Biogenic Volatile Organic Compounds (BVOCs)	111
5.1.1	Isoprene, Monoterpenes	111
5.1.2	Biomass Burning	112
5.1.3	MEGAN	112
5.2	Satellite HCHO measurements	112

5.2.1	Satellite Retrievals	112
5.2.2	OMI Algorithm BOAS	113
5.2.3	Optical Depth ( $\tau$ )	114
5.2.4	Scattering	115
5.2.5	Absorption cross section and number density	115
5.2.6	Air Mass Factors	115
5.2.7	OMI HCHO data products	116
5.2.8	HCHO Vertical Column Calculation	116
5.2.9	Uncertainty in OMI total columns (Moved)	119
5.2.10	Reference sector correction for comparison of products to various models	119
5.3	Recalculating HCHO from satellite(OMI) data over Australia	120
5.3.1	Process Outline	120
5.3.2	Quality filtering OMI HCHO slant columns	121
5.3.3	Reading OMHCHO daily slant columns	123
5.3.4	Regridding to 0.25 by 0.3125 8-day averaged vertical columns	123
5.3.5	Filtering pyrogenic HCHO	124
5.3.6	Filtering anthropogenic HCHO	126
5.3.7	Recalculating the AMF to create our own vertical HCHO columns	126
5.3.8	AMF code from Paul Palmer	130
5.3.9	Determination and application of the pacific ocean reference sector normalisation	130
5.3.10	Estimation of error or uncertainty (moved to Bioisop)	133
5.4	Validation and comparisons	133
5.4.1	Comparison with standard OMI product	133
5.4.2	Comparison with in-situ measurements	133
5.4.3	Summary	133
5.4.4	Conclusions	133
<b>A</b>	<b>Supplementary Notes</b>	<b>135</b>
A.1	Chemistry	135
	SOA	135
A.1.1	Relationship to Glyoxyl TODO: remove if never used	136
A.2	CAABA/MECCA	139
<b>B</b>	<b>Appendix A</b>	<b>141</b>
B.1	Grid Resolution	141
<b>C</b>	<b>Frequently Asked Questions</b>	<b>143</b>
C.1	How do I change the colors of links?	143

# List of Figures

1.1	Pressure (red) logarithmically decreasing, shown with percentage of atmosphere below at several points. Temperature (green) changes throughout the atmosphere. Figure edited from <a href="https://climate.ncsu.edu/Structure">https://climate.ncsu.edu/Structure</a> . . . . .	2
1.2	Tropospheric ozone processes, Figure 1 in <b>Young2018</b> DOI: <a href="https://doi.org/10.1525/element">https://doi.org/10.1525/element</a> . . . . .	5
1.3	Ozone production rate dependent on NO <sub>x</sub> and VOC concentrations ( <b>Mazzuca2016</b> ). . . . .	8
1.4	The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from <b>IPCC_Chapter2</b> found at <a href="https://www.ipcc.ch/publications_and_data/ar4/wg1/en/faq-2-1.html">https://www.ipcc.ch/publications_and_data/ar4/wg1/en/faq-2-1.html</a> . . .	15
1.5	The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from <b>IPCC_AR5_WG1</b> chapter 8. . .	16
1.6	HCHO spectrum, with a typical band of wavelengths used for DOAS path measurements. This is a portion of an image from <b>Davenport2015</b> .	18
1.7	An example spectrum showing interferences used for species concentration measurements by GOME-2. Image by EUMETSAT and ESA ( <b>GOME2Image</b> ). . . . .	19
1.8	Standard box model parameters, image taken from <b>Jacob_1999_book</b> .	21
1.9	Part of a figure from <b>Guenther2006</b> showing global isoprene emission factors. . . . .	26
2.1	MEGAN schematic, copied from <b>Megan_Website</b> . . . . .	33
2.2	Surface HCHO simulated by GEOS-Chem with UCX (top left), and without UCX (top right), along with their absolute and relative differences(bottom left, right respectively). Amounts simulated by GEOS-Chem for the 1st of January, 2005. . . . .	37
2.3	As figure 2.2, except looking at isoprene. . . . .	38
2.4	As figure 2.2, except looking at ozone. . . . .	39
2.5	Image from <b>Lee2015</b> . . . . .	41

2.6	Figure 1 and Table 1 from <b>Schenkeveld2017</b> with the following caption “An impression of OMI flying over the Earth. The spectrum of a ground pixel is projected on the wavelength dimension of the charge-coupled device (CCD; the columns). The cross-track ground pixels are projected on the swath dimension of the CCD (the rows). The forward speed of $7 \text{ kms}^{-1}$ and an exposure time of 2 s lead to a ground pixel size of 13 km in the flight direction. The viewing angle of $114^\circ$ leads to a swath width on the ground of 2600 km.” The table shows the optical properties for OMIs three channels. . . . .	43
2.7	Top panel: surface temperature averaged over January 2005. Bottom panel: surface temperature correlated against temperature over January 2005, with different colours for each gridbox, and the combined correlation. A reduced major axis regression is used within each gridbox (shown in top panel) using daily overpass time surface temperature and HCHO amounts (ppbv). The distribution of slopes and regression correlation coefficients (one datapoint per gridbox) for the exponential regression is shown in the embedded plot. . . . .	49
2.8	As figure 2.7 but for northern Australia. . . . .	50
2.9	As figure 2.7 but for south-western Australia. . . . .	51
2.10	GEOS-Chem mid-day tropospheric column $\text{NO}_2$ vs OMNO2d columns (averaged to match GEOS-Chems lower resolution). Absolute and relative differences, along with correlation shown on bottom row. . . . .	52
2.11	Top row (left to right): GEOS-Chem $\text{NO}_2$ mid-day tropospheric columns, OMNO2d $\text{NO}_2$ columns, modelled anthropogenic NO emissions. Second row: absolute and relative difference between GEOS-Chem and OMI $\text{NO}_2$ data, and the correlation. Third row: correlation between GEOS-Chem tropospheric column $\text{NO}_2$ and emitted NO, then between the model-satellite bias and the emissions. All correlation plots are coloured by emission rates. . . . .	54
2.12	As figure 2.11, for Autumn 2005. . . . .	55
2.13	As figure 2.11, for Winter 2005. . . . .	56
2.14	As figure 2.11, for Spring 2005. . . . .	57
2.15	As figure 2.11, except anthropogenic NO emissions are replaced by soil NO emissions. . . . .	58
2.16	As figure 2.11, for Autumn 2005, with soil NO emissions replacing anthropogenic NO emissions. . . . .	59
2.17	As figure 2.11, for Winter 2005, with soil NO emissions replacing anthropogenic NO emissions. . . . .	60
2.18	As figure 2.11, for Spring 2005, with soil NO emissions replacing anthropogenic NO emissions. . . . .	61
3.1	Top panel: isoprene emissions for January, 2005, shown in red, copletted with tropospheric hcho columns, shown in magenta. Both series are daily averages over Australia. Bottom panel: (RMA) linear regressions from between emissions of isoprene and tropospheric hcho columns, sampled randomly from the $2^\circ$ by $2.5^\circ$ latitude longitude gridboxes over Australia for the month of January (2005). . . . .	71

3.2	Top row is isoprene emissions for the month of January, in 2005, from GEOS-Chem and estimated from OMI respectively. Bottom row shows the absolute and relative differences between the two. . . . .	72
3.1	Ozonesonde release sites and the regions used to examine STT effect on tropospheric ozone levels. . . . .	65
3.2	Multi-year monthly median tropopause altitude (using the ozone defined tropopause) determined from ozonesondes measurements at Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013) (solid lines). Dashed lines show the 10th to the 90th percentile of tropopause altitude for each site. . . . .	67
3.3	Multi-year mean seasonal cycle of ozone mixing ratio over Davis, Macquarie Island, and Melbourne as measured by ozonesondes. Measurements were interpolated to every 100 m and then binned monthly. Black and red solid lines show median ozone and lapse-rate defined tropopause altitudes (respectively), as defined in the text. . . . .	68
3.4	An example of the STT identification and flux estimation methods used in this work. The left panel shows an ozone profile from Melbourne on 8 January 2004 from 2 km to the tropopause (blue dashed horizontal line). The right panel shows the perturbation profile created from band-pass filtering of the mixing ratio profile. The STT occurrence threshold calculated from the 95th percentile of all perturbation profiles is shown as the orange dashed line, and the vertical extent of the event is shown with the purple dashed lines (see details in text). The ozone flux associated with the STT event is calculated using the area outlined with the orange dashed line in the left panel. . . . .	70
3.5	Seasonal cycle of STT event frequency at Davis (top), Macquarie Island (middle), and Melbourne (bottom). Events are categorised by associated meteorological conditions as described in the text, with low pressure fronts ("frontal") in dark blue, cut-off low pressure systems ("cut-off") in teal, and indeterminate meteorology ("misc") in cyan. Events that may have been influenced by transported smoke plumes are shown in red (see text for details). . . . .	74
3.6	Seasonal distribution of STT events using the alternative STT proxy, obtained from consideration of the static stability at the ozone and lapse rate tropopauses, for Davis (2006-2013), Macquarie Island (2004-2013), and Melbourne (2004-2013). . . . .	75
3.7	The distribution of STT events' altitudes at Davis (top), Macquarie Island (middle), and Melbourne (bottom), determined as described in the text. Events are coloured as described in Fig. 3.5. . . . .	76
3.8	The distribution of STT events' depths, defined as the distance from the event to the tropopause, at Davis (top), Macquarie Island (middle), and Melbourne (bottom), determined as described in the text. Events are coloured as described in Fig. 3.5. . . . .	77

3.9	Comparison between observed (black) and simulated (pink, red) tropospheric ozone columns ( $\Omega_{O_3}$ , in molecules $\text{cm}^{-2}$ ) from 1 January 2004 to 30 April 2013. For the model, daily output is shown in pink, while output from days with ozonesonde measurements are shown in red. For each site, the model has been sampled in the relevant grid square. . . . .	78
3.10	Observed and simulated tropospheric ozone profiles over Davis, Macquarie Island, and Melbourne, averaged seasonally. Model medians (2005-2013 average) are shown as red solid lines, with red dashed lines showing the 10th and 90th percentiles. Ozonesonde medians (over each season, for all years) are shown as black solid lines, with coloured shaded areas showing the 10th and 90th percentiles. The horizontal dashed lines show the median tropopause heights from the model (red) and the observations (black). . . . .	79
3.11	Example comparisons of ozone profiles from ozonesondes (black) and GEOS-Chem (red) from three different dates during which STT events were detected from the measurements. The dates were picked based on subjective visual analysis. The examples show the best match between model and observations for each site. GEOS-Chem and ozonesonde pressure levels are marked with red and black dashes respectively. . . . .	80
3.12	Top panel: tropospheric ozone attributed to STT events. Bottom panel: percent of total tropospheric column ozone attributed to STT events. Boxes show the inter-quartile range (IQR), with the centre line being the median, whiskers show the minimum and maximum, circles show values which lie more than 1.5 IQR from the median. Values calculated from ozonesonde measurements as described in the text. . . . .	83
3.13	(Top) Tropospheric ozone, ( <i>I</i> )mpact per event, and ( <i>P</i> )robability of event detection per sonde launch, averaged over the region above Davis. The tropospheric ozone column $\Omega_{O_3}$ (black, left axis) is from GEOS-Chem, while the STT probability <i>P</i> (magenta, right axis) and impact <i>I</i> (teal, right axis) are from the ozonesonde measurements. The STT impact is multiplied by ten to better show the seasonality. (Bottom) Estimated contribution of STT to tropospheric ozone columns over the region, with uncertainty (shaded area) estimated as outlined in Sect. 3.6. The black line shows STT ozone flux if event lifetime is assumed to be two days, with dashed lines showing the range of flux estimation if we assumed events lasted from one day to one week. . . . .	84
3.14	As described in 3.13, for the region containing Macquarie Island. . . . .	85
3.15	As described in 3.13, for the region containing Melbourne. . . . .	86
5.1	OMI uncertainty before and after gridding and averaging 8 days from Jan 1 2005 to Jan 8 2005. The third panel shows the number of pixels in each grid box after 8 days of averaging, before accounting for fire. . . . .	113
5.2	Solar and viewing zenith angles, image copied from <b>SZA_Image</b> originally from a NASA website. . . . .	117

5.3	Column density histograms for a subset of OMI swaths over Australia on the 18th of March 2013. Negative entries are shown in the left panel, positive in the right, note the different scale between negative and positive panels. . . . .	123
5.4	Example of grid space change using 0.5x0.5 and 0.25x0.3125 latitude by longitude resolution. . . . .	125
5.5	Example of MODIS 8 day grid interpolation from 0.5x0.5 to 0.25x0.3125 latitude by longitude resolution. This example uses MODIS fire counts for 1-8 January 2005. . . . .	126
5.6	Vertical column HCHO calculated using OMI satellite swaths with GEOS-Chem aprioris, averaged over 1-8 January 2005 with and without fire affected squares removed. . . . .	127
5.7	Constructed example of the initial interpolation of OMI's $\omega$ onto a pressure dimension with mismatched surface pressure. . . . .	129
5.8	Example of remote pacific reference sector correction (RSC) using 8-day average measurements and one month modelled data. $\Omega_{VC}$ shows the uncorrected vertical columns, while $\Omega_{VCC}$ shows the corrected vertical columns. OMI corrections shows the correction applied globally based on latitude and OMI track number(sensor). $\Omega_{GC}$ shows the GEOS-Chem modelled HCHO VC over the RSC, with $\Omega_{VCC}$ showing the corrected VC over the same area. . . . .	131
5.9	Example of track correction interpolations for January 1st 2005, points represent satellite slant column measurements, with lines interpolating and extrapolating along the latitudinal direction. . . . .	132
A.1	The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from <b>IPCC_Chapter2</b> found at <a href="https://www.ipcc.ch/publications_and_data/ar4/wg1/en/faq-2-1.html">https://www.ipcc.ch/publications_and_data/ar4/wg1/en/faq-2-1.html</a> . .	137
A.2	The overall radiative forcings and uncertainties of several atmospheric constituents This is an image taken from <b>IPCC_AR5_WG1</b> chapter 8. . .	138





# List of Tables

3.1	HCHO yields from various species averaged over Australia during Summer. . . . .	74
3.2	HCHO yields from various species, and lifetime against oxidation by OH. . . . .	75
3.1	Number of sonde releases at each site over the period of analysis. . . . .	66
3.2	Total number of ozonesonde detected STT events, along with the number of events in each category (see text). . . . .	73
3.3	Seasonal STT ozone contribution in the regions near each site, in $\text{kg km}^{-2} \text{ month}^{-1}$ . In parentheses are the relative uncertainties. . . . .	82
5.1	OMI quality flag values table from <b>Kurosu2014</b> . . . . .	122
B.1	The 47 level vertical grid used by GEOS-Chem . . . . .	142





## Chapter 3

# Biogenic Isoprene emissions in Australia

### 3.1 Introduction

Isoprene has a large impact on the oxidative properties of the atmosphere, as it reacts quickly with the OH radical to form RO<sub>2</sub>. These react quickly with NO<sub>x</sub> to form OVOCs (such as HCHO), SOAs, and ozone. Australian isoprene emissions are poorly understood due to poor measurement coverage, and poor emission factor characterisation. The emissions of isoprene have been modelled at around 500 Tg C yr<sup>-1</sup> in **Guenther1995**; **Guenther2006** using MEGAN, and more recently around 465 Tg C yr<sup>-1</sup> in **Messina2016** using ORCHIDEE. The global emission factors model used to derive these estimates is based on modelling emissions from different plant species (phenotypes), and very few are used to set the emission factors of Australian forests.

We estimate isoprene emissions in Australia using top-down estimates based on recalculated OMI HCHO measurements and modelled isoprene to HCHO yields. These estimates are compared to several campaigns (SPS1, SPS2, MUMBA, Daintree) and used as the new boundary conditions for GEOS-Chem. Sensitivity to soil moisture, (maybe) LAI, and satellite AMF calculation is examined and quantified for some scenarios. The effect of using these new top down isoprene emissions as the boundary conditions for GEOS-Chem is studied. Wellness of fit between in-situ (at Wollongong) HCHO, satellite (OMI), and modelled (GEOS-Chem) HCHO is determined with and without updated emissions estimates.

One of the most popular emissions inventories for biogenic isoprene, the Model of Emissions of Gases and Aerosols from Nature (MEGAN). Global atmospheric studies often use MEGAN along with a chemical transport model (CTM) to examine transport, deposition, and various chemical processes in the atmosphere. Emissions of Biogenic Volatile Organic Compounds (BVOCs) including isoprene are often the subject of studies as they are still relatively uncertain, as well as being drivers for important oxidation and pollution events.

(MEGAN) is poorly calibrated for Australian conditions, their emissions of isoprene (C<sub>5</sub>H<sub>8</sub>) may be overestimated, especially in the southeast. **Muller2008** compared MEGAN against emissions calculated using top down estimates from the GOME2 satellite measurements of formaldehyde. **Stavrakou2015** showed that this overestimate may be a factor of 2-3 in January. **Sindelarova2014** show how 50% of the isoprene emissions could be reduced by accounting for lower soil moisture. **Emmerson2016**

discuss the suitability of MEGAN's isoprene and monoterpene emission factors over southeast Australia, and suggest isoprene emissions are estimated 2-6 times too high. They also show that no blanket increase or decrease in emission factors is appropriate for the entire southeast of Australia.

In the remote troposphere HCHO production is dominated by methane oxidation, while in the continental boundary layer (CBL) production is largely due to NMVOCs (Abbot2003; Kefauver2014). This suggests that HCHO enhancement over continents can be used to determine NMVOC emissions. In the CBL, HCHO enhancement is generally driven by short lived ( $< 1$  hr) precursors (most importantly isoprene). HCHO itself has a lifetime of a few hours (Kefauver2014). Isoprene is emitted and enters the atmosphere in the gas phase, where it begins a complex series of reactions. Formaldehyde is produced with high yields in many of the isoprene reactions, which are discussed in more detail in Section 1.3.3. HCHO measurements are often used as a check on how well isoprene reactions are simulated, as model output can then be compared against them (Marvin2017).

Isoprene and monoterpene emissions are both very uncertain in Australia. Monoterpene oxidation by  $O_3$ , OH and  $NO_3$  radicals may also form aerosols, with the reaction with ozone forming the most particles (Kanakidou2005).

### 3.1.1 Top-down emissions estimates

There are now a few methods of estimating isoprene emissions using satellite measurements of emission products, here I describe them and briefly compare the pros and cons of each.

#### Linear

This technique is the simplest, and is performed in this thesis. With the vertical columns of biogenic HCHO we can infer the local (grid space) isoprene emissions using effective molar formaldehyde yield (In other continents around 2-3, or 1 in low  $NO_x$  conditions) (Palmer2003; Marais2012; Bauwens2016). If we assume there is fast HCHO yield, so that the effect of chemical transport is minimal, and that HCHO and isoprene are at steady states, then we can calculate local yield from our CTM. Yield is calculated from the modelled slope between isoprene emissions and HCHO total column within each gridbox over Australia, as performed in Palmer2003 using modelled values between 1200-1400 LT which is around the overpass time of the OMI. This modelled yield is then used in conjunction with the recalculated OMI measurements in order to estimate isoprene emissions. To calculate emissions we use a reduced major axis (RMA) regression between modelled average values of the loss rates and total columns, an example is shown in figure TODO: figure with RMA of these over whatever time and space I end up using.

#### Bayesian

Satellite based emissions estimates may allow us to improve the models without requiring lots of hard work on calibrating MEGAN to the large data sparse continent of Australia. Emissions of monoterpenes ( $C_{10}H_{16}$ , two units of isoprene) may also be

underestimated in southeastern Australia, which could lead to the unique scenario of neither type of emission dominating VOC chemistry over the forests (**Emmerson2016**).

Another method of correcting isoprene emissions using observed HCHO total column involves a Bayesian inversion. **Shim2005** work with GOME HCHO observations and GEOS-Chem, looking at areas with high signal to noise ratio (higher HCHO concentrations). They show that the model underestimates isoprene emissions and HCHO concentrations by 14-46%, with the corrected VOC emissions reducing the model biases to 3-25%.

The Bayesian inversion is also used in **Curci2010** where a regional CTM (CHIMERE) simulates HCHO, which is compared against OMI observed HCHO and shown to be regionally biased. This bias is expected to be caused by errors in MEGANs isoprene emissions estimations. The CHIMERE model is used to derive yields of HCHO from the various local VOCs and these are then used in estimating local emissions. The model is run initially with emissions of BVOCs and reactive anthropogenic VOCs (RAVOCs) turned off in order to work out the background (b) values of these compounds. The Bayesian inversion is used to correct regionally biased biogenic isoprene emissions by optimising these parameters in order to simulate HCHO closest to the observed HCHO levels. **Curci2010** uses CHIMERE as the forward model to determine the relationship between HCHO ( $y$ ), isoprene and reactive anthropogenic VOCs ( $x$ ), using

$$y = Kx + b + \epsilon \quad (3.1)$$

where  $\epsilon$  are the (assumed) independent errors in measurements.  $K$  is the Jacobian matrix determined from CHIMERE representing the sensitivity of  $y$  to the state variable  $x$ . This  $K$  matrix is used in conjunction with error covariance in  $x$  to determine the Maximum A Posteriori (MAP) solution to calculate the optimal estimate of  $x$  ( $\hat{x}$ ).

### 3.1.2 Aims

Here we introduce how uncertain isoprene emissions are over Australia, and discuss literature which shows how the estimates may be too high. Section ?? describes the model, satellite, and campaign data we use to determine and analyse isoprene emissions. The OMI measurements used in this research are recalculated using an updated estimate of HCHO profiles and validated against Wollongong total column measurements. Section 3.2 lays out how the isoprene emissions are estimated, and the results are examined in Section 3.3

## 3.2 Methods

### 3.2.1 Outline

Here is an overview of the steps involved in my Thesis, which take satellite data and model output to estimate isoprene emissions.

1. Download Aqua/Terra MODIS gridded fire counts (MOD14A1), smoke measurements (OMAERUVd), and Aura HCHO columns (OMHCHO). These products are discussed in more detail in Section ??.

2. Run GEOS-Chem with satellite overpass output averaged over 1200-1300 LT (matching Aura overpass time).
3. Recalculate OMHCHO vertical columns using GEOS-Chem to recreate the shape factors for each slant column. Also done with code from Paul Palmers group which includes recalculation of the scattering weights. The method to read the satellite data is given in Section ??, while reprocessing the column AMFs is detailed in Section ??.
4. Mask HCHO columns which occur on days with non-zero (MODIS) fire counts over the prior 8 days, as is done in **Marais2012** More detail on this step can be found in ??
5. TODO: Filter gridsquares where OMI AAOD (OMAERUVd) is over some threshold to be determined (ALSO TODO)
6. TODO: Filter NO<sub>x</sub> influenced grid squares.
7. Biogenic Yield calculated in each 2x2.5 ° grid square using daily averaged HCHO from biogenic only run of GEOS-Chem, linearly correlated with MEGAN isoprene emissions within those squares
8. TODO: Biogenic yield masked when smearing term is above a threshold (TODO: determine)
9. Yield (actually regression slope between hcho and emissions) multiplied against recalculated OMI HCHO to provide top down isop emissions estimate
10. Compare against MEGAN, MUMBA, SPS, Daintree, Wollongong, and one set of airplane(?) measurements

### 3.2.2

### 3.2.3 Satellite inversion

Top-down estimates look at how much of a chemical is in the atmosphere and try to work out how much of its major precursors were emitted. This generally takes advantage of longer lived products which may reach a measurable equilibrium in the atmosphere. For isoprene this is done by looking at atmospheric HCHO enhancement, which can be largely attributed to isoprene emissions once transport and other factors are accounted for. Recently **Stavrakou2015** used satellite HCHO measurements to constrain anthropogenic sources of isoprene and found good global agreement with the bottom up estimates, although some regions had sources differ by up to 25-40%. Their study used the RETRO 2000 database for anthropogenic emission aprioris except for Asia in 2008 where REASv2 was used. Since 1997, when GOME first measured HCHO over Asia (**Thomas1998**), satellites have been able to provide a total column measurement of HCHO, one of the primary products of isoprene.

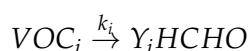
Satellites recording reflected solar spectra use DOAS to measure various trace gases in the atmosphere, including formaldehyde. While satellite measurements can only be used during daytime hours, HCHO lifetimes are sufficiently short that any

night-time chemistry will not affect midday observations (**Wolfe2016**). Satellites can be used to measure the seasonal and interannual variability of HCHO over the globe. These records can be compared with modeled estimates of HCHO and used as a proxy to estimate isoprene emissions. This has been done in North America (**Palmer2003; Millet2006**), South America, Africa, China, Europe (**Dufour2009**), and recently globally (**FortemsCheiney2012; Bauwens2016**). Often these works use two forms of measurement such as satellite and aircraft data combined for validation (**Marais2014**). There is less information available from satellite measurements at higher latitudes due to increased errors (**DeSmedt2015**).

Using HCHO to determine emissions of isoprene was initially performed by **Palmer2001; Palmer2003** who used in-situ summertime HCHO measurements over North America as model validation. Isoprene emissions fluxes were derived using the Global Ozone Monitoring Experiment (GOME) satellite instrument. Palmer's method improved biogenic isoprene emissions estimates (compared with in-situ measurements) over two available inventories: the U.S. EPA Biogenic Emissions Inventory System (BEIS2) and the Global Emissions Inventory Activity (GEIA). This showed an inversion technique which could be used to improve large scale emissions estimates without further expensive measurement campaigns.

TODO: Read through this list of sources on the hcho to isop process : taken from Wolfe2015 Such techniques have informed isoprene emission inventories in North America (Abbot et al., 2003; Millet et al., 2008 (**Palmer2003; Millet2006; Palmer2006**)), South America (**Barkley2013**), 2008), Europe (**Curci2010; Dufour2009**), Africa (**Marais2012**), Asia (Fu et al., 2007; Stavrou et al., 2014), and globally (Fortems-Cheiney et al., 2012; (**Shim2005**); Stavrou et al., 2009).

Initially studies assumed a simple linear steady-state relationship between HCHO and its precursors (**Palmer2003; Palmer2006; Millet2006**). This allowed a simple calculation of isoprene using the measured HCHO, with estimated reaction rates and yields. The methodology for calculating VOCs from HCHO is laid out in **Palmer2003** and takes into account the expected lifetime and reaction rates of the precursor VOCs and HCHO. Assuming HCHO is produced quickly from short-lived intermediates, and the column is at steady state:



Where  $Y_i$  is HCHO yield per C atom (a measure of how much HCHO will form per gram of C from a VOC within a system), and  $k_i$  is the reaction rate. Then assuming a steady state of atmospheric HCHO ( $\Omega$  molecules  $cm^{-2}$ ) produced by oxidation of VOCs ( $VOC_i$ ) and no horizontal transport:

$$\Omega = \frac{1}{k_{HCHO}} \sum_i Y_i E_i$$

Where  $i$  indexes a chemical species,  $k_{HCHO}$  is the HCHO loss rate due to OH and photolysis,  $Y_i$  is the molar HCHO yield from oxidation of  $i$ , and  $E_i$  is emission fluxes (C atoms  $cm^{-2}s^{-1}$ ).

Estimates of  $Y_i$  can be attained from a model as shown in **Millet2006** This involves a reduced major axis (RMA) correlation calculation between modelled HCHO and



isoprene columns, multiplied by their loss rates (to photolysis and oxidation) (as a normalising factor). In high NO<sub>x</sub> environments where HCHO has a lifetime on the order of 30 minutes, it can be used to map isoprene emissions with spatial resolution from 10-100 kms. Horizontal transport 'smears' the HCHO signal so that source location would need to be calculated using windspeeds and loss rates (**Palmer2001**; **Palmer2003**). Smearing is explicitly handled in these studies due to the importance of transport and NO<sub>x</sub> on forming robust and accurate estimates. Over Australia NO<sub>x</sub> levels are generally not high enough to ensure quick HCHO formation and we must take extra care that we can account for the transport or 'smearing' caused by slower HCHO formation, details on this process can be found in Section 3.2.7.

More recently, full inversions that better account for transport, source attribution, and chemical schemes have been implemented (**FortemsCheiney2012**). TODO: full description of this better inversion technique going through FortemsCheiney2012.

**Kefauver2014** reviews remote sensing of BVOCs, which are on the rise, examining the last 20 years of data and analysis of the satellite products. Their review encompasses the latest reports up to 2014. The modelled isoprene and BVOC emissions from MEGAN (**Guenther2000**) of 500 and 1150 Tg<sup>a</sup> respectively are still the global go to estimates. Their review reinforces the message that NMVOCs affect the oxidative capacity of the atmosphere and are largely driven by and sensitive to vegetation. The tropospheric affects from NMVOCs on the hydroxyl radical (OH), ozone (O<sub>3</sub>), SOAs, and methane longevity, all interconnect to form a very complex system which still suffers from relatively large uncertainties in both measurement and chemistry mechanisms. One focus of **Kefauver2014** is HCHO, which is the dominant product of most BVOCs which is measurable by remote sensing. The main datasets of HCHO are from four satellite instruments: GOME on ERS-2, SCIAMACHY on ENVI-SAT, OMI on EOS AURA, and GOME2 on MetOp-A. These satellites have slightly different spectral and spatial resolutions, as well as using varied processes to estimate HCHO from detected radiances. This can lead to different estimates between instruments or methodologies as described in **Lorent2017** which means validation and comparison is more important when using these remotely sensed data.

Total HCHO is measured by satellite over the entire world, however the technique is not perfect and suffers from uncertainties and interferences. Satellite based chemical concentrations rely on ground-based measurements and modelled data for validation. They provide various readings with daily global coverage which is not otherwise feasible.

Validation is important due to the various uncertainties in the satellite remote sensing process, with apriori assumptions having the greatest effect on structural uncertainty between measurements techniques **Lorent2017** **Zhu2016** use SEAC<sup>4</sup>RS aircraft HCHO measurements over the southeastern US as model validation, and show a bias in the assumed OMI shape factor that leads to a bias between satellite and SEAC<sup>4</sup>RS measurements. **Marais2014** compare OMI based isoprene emission estimates against relaxed eddy accumulation measurements from African field campaigns, as well as MEGAN and GEOS-5 inventories. **Dufour2009** use HCHO from SCIAMACHY, and examine Europe using CHIMERE as the chemical model. In their work they show that satellite measurements can reduce source emission uncertainty by a factor of two, where emissions are relatively large.

### 3.2.4 Calculation of Emissions

As is done in **Palmer2003; Millet2006; Bauwens2016** we assume that HCHO, and Isoprene columns are in a steady state, with no horizontal transport. We also assume that isoprene is the only compound enhancing the HCHO levels, which requires that we filter out influence from fires. Emissions of precursors are easy to calculate as long as we know the molar HCHO yields ( $Y_i$ ) and effective chemical loss rates ( $k_i$ ):

$$\Omega_{HCHO} = \frac{1}{k_{HCHO}} \sum_i k_i Y_i \Omega_i = \frac{1}{k_{HCHO}} \sum_i Y_i E_i \quad (3.2)$$

This works if there is fast HCHO yield, so that the effect of chemical transport is minimal. The background HCHO is calculated using measurements in the remote pacific at the same time and latitude. Table 3.1 shows the average yield calculated for Australia. (TODO: this table and some notes)

In order to approximate the isoprene to HCHO yields over Australia, GEOS-Chem is run and the slope ( $S$ ) between modelled tropospheric HCHO columns and emissions of isoprene within each grid box. Figure (TODO: example grid box and regression plot.) shows the regression between emitted isoprene and tropospheric column HCHO, averaged between 1200-1300 LT each day. We can infer the local (grid space) isoprene emissions ( $E_{isop}$ ) using effective formaldehyde yield from isoprene ( $Y_{isop}$ ).

$$\Omega_{HCHO} = S \times E_{isop} + B \quad (3.3)$$

Where  $B$  is the background HCHO, and  $S = Y_{isop}/k_{HCHO}$  is determined monthly as the regression between  $\Omega_{HCHO}$  and  $E_{isop}$  on daily saved outputs from GEOS-Chem over Australia using 2 by 2.5° horizontal resolution. Modeled background emissions can be ignored here as they do not affect the slope calculation. Once we have calculated this slope, we use the same formula (Eqn. 3.3) to determine the isoprene emissions. By replacing  $\Omega_{HCHO}$  and  $B$  with OMI based values,  $E_{isop}$  is the only unknown.

In several studies OMI satellite HCHO columns are scaled up by up to 40% to match in-situ measurements TODO:citations. Since we don't have enough in-situ data to reasonably scale the satellite columns we instead re-run the calculation with them scaled up by 40% and consider this as a possible bound on satellite uncertainty. Yield and estimated top-down emissions are therefore given an upper and lower bound on satellite based uncertainty through this method.

There are a couple of ways to determine the modelled background HCHO concentration, one of which involves running the model with isoprene emissions turned off, which allows us to see exactly how much the modelled isoprene emissions alter each vertical column of HCHO. This is effective since we have assumed variation in HCHO columns only depends on isoprene emissions, so our background term is theoretically identical to the emission free simulated HCHO. The other way involves looking at HCHO over the remote pacific at matching latitudes and times, which emulates how the background is determined for the measured HCHO. These modelled background HCHO concentrations are mostly used for comparison with other datasets. TODO: show how these two numbers compare! figure with time series of both backgrounds for Aus, SEAus, and remote ocean.

The Background from OMI is determined using the mean column HCHO measured over the remote pacific ocean (180-120°W). TODO: Update the background term to do as follows: For this term we average each month of remote ocean measurements, as well as averaging longitudinally within 180-120°W, and finally  $B$  is estimated at each latitude using  $\pm 10^\circ$ . This gives us a background which is appropriate for any latitude, and is shown in Figure TODO: figure with background region highlighted and a time series of background values. When calculating the  $E_{isop}$  from our modeled slope with OMI HCHO and background, we end up with negative emissions whenever the OMI HCHO column is less than the OMI background (as  $E_{isop} = \frac{\Omega_{HCHO} - B}{S}$ ). These are set to zero, which increases the average by around TODO: X%. The measured background HCHO is the average concentration measured in the remote pacific at the same time.

Figure 3.1 shows the modelled isoprene emissions and column HCHO concentrations along with the RMA regression line, sampled from grid boxes over Australia for January 2005. Some affects from the low emissions in grid boxes which are largely oceanic can be seen and are handled by TODO: handle these and document here. Due to the low horizontal resolution of GEOS-Chem (2 by 2.5°, latitude by longitude), calculations from grid boxes on the coast which are largely oceanic need to be discarded as the change in HCHO is not dominated by emissions of isoprene, as is assumed for Eqn 3.3. A nested version of GEOS-Chem allows a much better analysis of coastal regions, at 0.25 by 0.3125° resolution.

Using this modelled slope at 2° by 2.5° and applying it to equation 3.3 with  $B$  and  $\Omega_{HCHO}$  calculated using OMI satellite measurements provides a new estimate of isoprene emissions. Figure 3.2 shows the emissions calculated this way along with the Emissions output by GEOS-Chem averaged over January, 2005.

### 3.2.5 Emissions drivers

Calculated yields of HCHO can be classified using a box model which approximates specific environments, as described in Section ?? . TODO: A table of different factors affecting emissions for three scenarios; urban, forest, shrublands is given in Table XX. The calculated yields for these scenarios is based on the CAABA/MECCA box model (described in Section ??) TODO: compare scenarios yields and show map of Australia with mapped closest scenario(one colour for each scenario, contourf).

### 3.2.6 HCHO Products and yield

Australian forests are strong emitters of both isoprene and monoterpenes, which go on to form various products including secondary organic aerosols, oxygenated VOCs (OVOCs), ozone, OH, and HO<sub>2</sub>. This production occurs over several steps, yields are often classed into at least two categories. First generation yield refers to the amount of HCHO produced per unit isoprene consumed by initial oxidation, total yield (sometimes molar yield) refers to time dependent yield of HCHO over multiple oxidation stages (Wolfe2016). Wolfe2016 define prompt yield as the change in formaldehyde measurement per unit change in initial isoprene emissions. Some argue that isoprene emissions are overestimated, due to the fact that they are based on relatively few measurements of isoprene emission factors (Winters2009; FortemsCheiney2012)

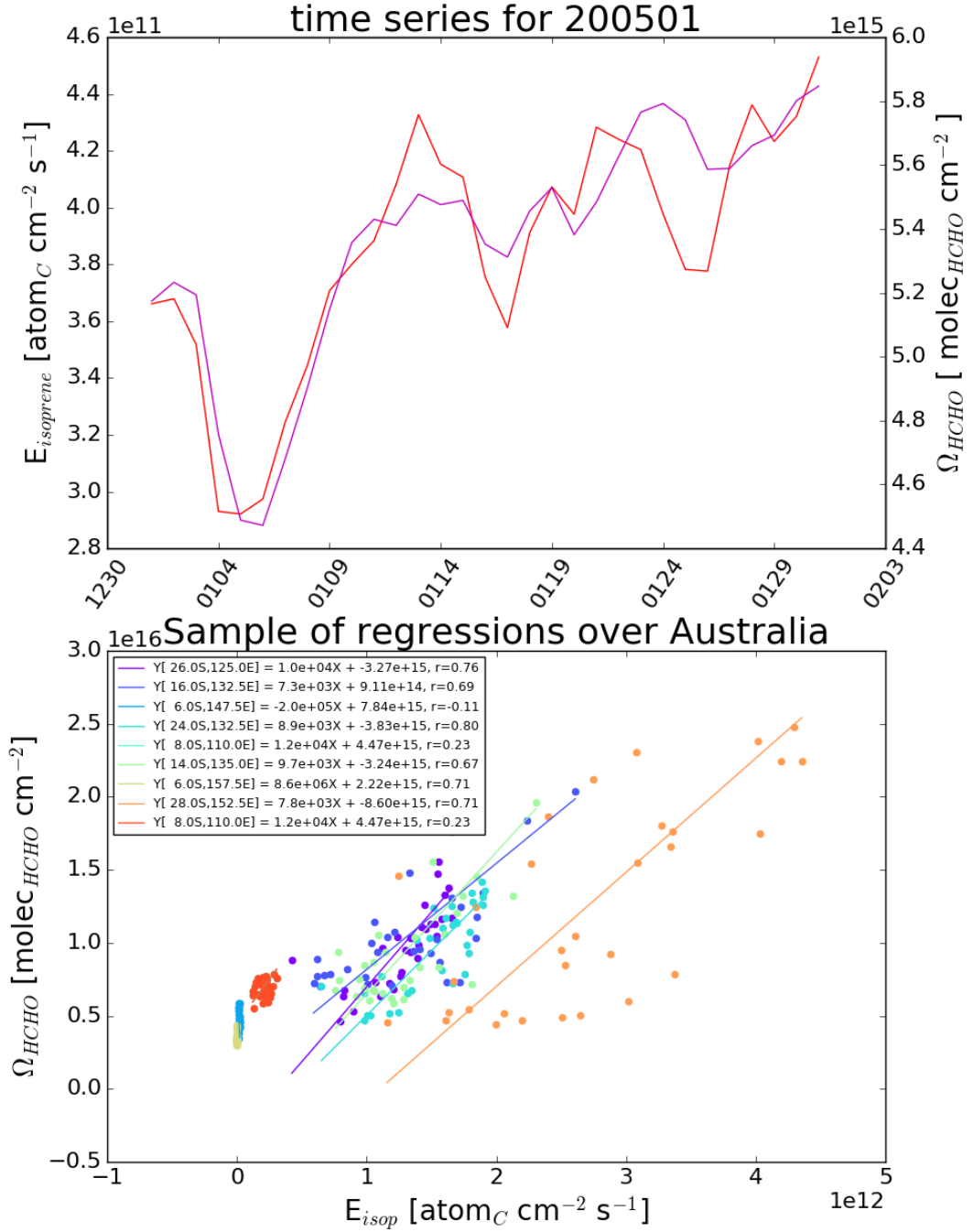


FIGURE 3.1: Top panel: isoprene emissions for January, 2005, shown in red, coplotted with tropospheric hcho columns, shown in magenta. Both series are daily averages over Australia. Bottom panel: (RMA) linear regressions from between emissions of isoprene and tropospheric hcho columns, sampled randomly from the  $2^\circ$  by  $2.5^\circ$  latitude longitude gridboxes over Australia for the month of January (2005).

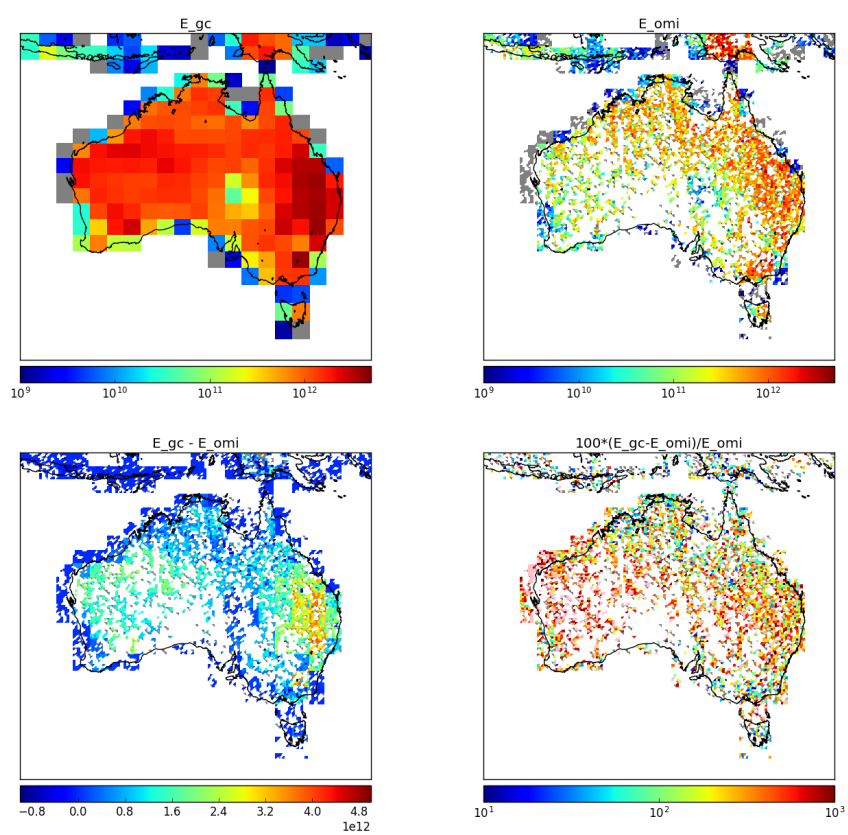


FIGURE 3.2: Top row is isoprene emissions for the month of January, in 2005, from GEOS-Chem and estimated from OMI respectively. Bottom row shows the absolute and relative differences between the two.

TODO: read and cite paper mentioned in Fortems. Recently **Emmerson2017** showed that MEGAN estimates 3-6 times too much isoprene emissions, and 4 times too little monoterpenes when compared against 4 (relatively small scale) measurement campaigns in southeastern Australia.

Isoprene production of HCHO depends on several factors, importantly  $\text{NO}_x$  levels have a direct effect on the fate of VOCs in the atmosphere. At higher NO mixing ratios (at least a few hundred pptv), organic peroxy radicals ( $\text{RO}_2$ ) react mostly with NO. At low NO (less than 10's of pptv), reaction with  $\text{HO}_2$ , other  $\text{RO}_2$ , and isomerization dominate the fate of  $\text{RO}_2$ . In low  $\text{NO}_x$  environments, reported HCHO yields from isoprene are from X to Y%, while in high  $\text{NO}_x$  environments this value is X to Y% TODO: these values from table. For monoterpenes the yields are around X, Y% for low, high  $\text{NO}_x$  respectively. Emissions and yields for various species including some terpenes can be seen in table 3.1. **Wolfe2016** determine that going from  $\text{NO}_x = 0.1$  to 2.0 ppbv triples the prompt yield of HCHO, from 0.3 to 0.9 ppbv  $\text{ppbv}^{-1}$  due to isoprene, while the background HCHO doubles. They determine prompt yield as the change in HCHO per change in  $\text{ISOP}_0$ , using  $[\text{ISOP}]_0 = [\text{ISOP}] \exp(k_1[\text{OH}]t)$ ; where  $k_1$  is first order loss rate. This effectively relates HCHO abundance with isoprene emission strength.

$\text{NO}_2$  measured by OMNO2d gives us a daily mid-day measurement which we can compare to output from GEOS-Chem to determine how well the model does at simulating  $\text{NO}_2$ . This is also done in **Travis2016** as a way to examine model bias in ozone (potentially due to  $\text{NO}_2$  bias) over the USA.

Looking at Australian emissions from running GEOS-Chem and using yields provided by XYZ (TODO other table), we see that Australia may be more or less likely to do something TODO: this comparison sentence would be good to tie up tables and be copied to conclusions.

Conversions between HCHO per unit C yield and molar % yield from species X given by the equation  $Y_{\text{molar}\%} = 100 \times C_X \times Y_{\text{HCHO per unit C}}$ , where  $C_X$  is how many Carbon are within species X (5 for isoprene, 10 for monoterpenes, etc...). For instance a 200% molar yield of HCHO from isoprene implies 1 Mole of  $\text{C}_5\text{H}_8$  becomes 2 Mole HCHO which is a 0.4 HCHO per unit C yield.

TODO: Fill out this table

Yields table looking at literature provided yields of HCHO.

### 3.2.7 Accounting for smearing

Accounting for transport of the precursors is important, especially in low  $\text{NO}_x$  conditions in which isoprene has a longer lifetime (days). When estimating emissions of isoprene using one of its products, it is often assumed that isoprene has a short lifetime, however when low  $\text{NO}_x$  environments (which are prevalent in the Australian outback) this assumption can be wrong. Smearing (or spatial smearing) is a measure of how much formaldehyde (the product) was created from isoprene (the precursor) emissions in a different grid box. Smearing has been measured in order to account for this uncertainty in various works (**Martin2003**; **Palmer2003**; **Millet2006**; **Marais2012**; **Barkley2013**; **Zhu2014**; **Wolfe2016**), often implementing the method designed in **Palmer2003**

TABLE 3.1: HCHO yields from various species averaged over Australia during Summer.

Species	Emissions <sup>a</sup> (Tg C per month)	Lifetime <sup>b</sup>	HCHO Yield <sup>c</sup> (per C reacted)	HCHO production <sup>d</sup> %
Isoprene	Y	n minutes	0.x	10
$\alpha$ -Pinene	Y	n minutes	0.x	10
$\beta$ -Pinene	Y	n minutes	0.x	10
HCHO	Y	n minutes	1.0	10

a: Calculated using GEOS-Chem emissions over Australia in January 2005.

b:

c:

d: Production determined by dividing emission\*yield by the sum of all VOC emissions\*yields.

Horizontal transport complicates estimation of precursor emissions, as the smearing length scale which increases beyond our gridbox size. The smearing length scale; the distance travelled downwind ( $L_{d,i}$ ) by a precursor (i) before becoming HCHO can be estimated using:

$$L_{d,i} = \frac{U}{k_i - k_{HCHO}} \ln \left( \frac{k_i}{k_{HCHO}} \right)$$

where U is wind-speed. **Palmer2003** further define a smearing length scale:  $L_{s,i}$  as the distance downwind where a fraction  $(1 - 1/e)$  of the precursor is completely transformed into HCHO. This equation uses the initial VOC column concentration ( $[VOC]_0$ ) at the point of emission and mass balance equations as follows:

$$\frac{1}{k_{HCHO} - k_i} \left( k_{HCHO} \exp \left[ \frac{-k_i L_{s,i}}{U} \right] - k_i \exp \left[ \frac{-k_{HCHO} L_{s,i}}{U} \right] \right) = \frac{1}{e} \quad (3.4)$$

with limiting values  $L_{s,i} \rightarrow U/k_i$  for  $k_i \ll k_{HCHO}$ , and  $L_{s,i} \rightarrow U/k_{HCHO}$  for  $k_{HCHO} \ll k_i$ .

TODO: calculation of smearing Smearing sensitive grid boxes within the model can be detected by running the model with two times with only isoprene emissions changed. Similarly to smearing sensitivity calculations in **Marais2012** we run GEOS-Chem with isoprene emissions halved, then calculate  $\hat{S} = \frac{\Delta \Omega_{HCHO}}{\Delta E_{Isop}}$ , where  $\Delta$  represents the monthly mean departure over 1300-1400LT from default run values. This allows us to determine which gridboxes are disproportionately affected by emissions from non-local sources. Consider halving the isoprene emitted globally and rerunning the model, you would expect HCHO enhancement (above background levels) to be halved in isoprene emitting grid-squares. If the local grid-square HCHO enhancement is reduced by much more than half (factoring yield) then you can infer sensitivity to non-local isoprene emissions.

Smearing can be dependent on local or regional weather patterns, as greater wind speeds will reduce the time any emitted compound stays within the local grid box. As such smearing sensitivity is both spatially and temporally diverse, shown in figure TODO: is a picture of the smearing sensitivity over Australia. Large smearing



TABLE 3.2: HCHO yields from various species, and lifetime against oxidation by OH.

Species	HCHO Yield (molar % )	Life vs OH	NO <sub>x</sub> background	Source
Isoprene	315±50	35 min	High	a
	285±30		High	a
	225		High	b
	150		Low	b
	150		Low	d
	450		High	d
	235		1 ppbv	e
	150		0.1 ppbv	e
$\alpha$ -Pinene	28±3	1 hour	Low	c
	X±3		X	d
	230±90		High	a
	190±50		High	a
	19			b
	210		1 ppbv	e
	70		0.1 ppbv	e
$\beta$ -Pinene	65±6	40 min	Low	c
	X±3		X	d
	540±50		High	a
	450±80		High	a
	45			b
Methane	100	1 year		b
Ethane	180	10 days		b
Propane	60	2 days		b
Methylbutanol	.13(per C)	1 hour		b
HCHO	100	2 hour		b
Acetone	.67(per C)	10 days		b
Methanol	100	2 days		b

a **AtkinsonArey2003** Table 2, Yield from Isoprene reaction with OH, two values are from two referenced papers therein.

b **Palmer2003** lifetimes assume  $[\text{OH}]$  is  $1\text{e}15 \text{ mol cm}^{-3}$ .

c (**Lee2006**): Calculated through change in concentration of parent and product linear least squares regression. Estimates assume 20° C conditions.

d **Wolfe2016** “prompt yield”: change in HCHO per change in  $\text{ISOP}_0$ .  $[\text{ISOP}]_0 = [\text{ISOP}] \exp(k_1[\text{OH}]t)$ ; where  $k_1$  is first order loss rate. Effectively relates HCHO abundance with isoprene emission strength

e **Dufour2009** One-day yields from oxidation modelled by CHIMERE, using MCM reference scheme.

f Calculated using PTR-MS and iWAS on SENEX campaign data.



values can be seen near many coastlines as the Emissions are very low, which makes transported isoprene relatively more important in these gridboxes. Once the smearing sensitive grid squares are filtered out, application of equation ?? gives us an estimate of isoprene emissions across the nation.

Most recently a **Bauwens2016** undertook a similar process to what I am doing, although with slightly different focus, using the IMAGESv2 global CTM instead of GEOS-Chem. They calculate emissions which create the closest match between model and satellite vertical columns, and compare these postiori data with the apriori (satellite data) and independent data sets. (TODO: simple outline of what they did and how my focus is different, this paper will also need to be summarised in the LitReview) TODO: Plots of  $\hat{S}$  showing worst smearing affected areas per season.

### 3.3 Results

#### 3.3.1 Emissions comparisons

Some global numbers (TODO: where do I throw these?) **Guenther2012** Estimate global biogenic isoprene emissions at roughly  $535 \text{ Tg yr}^{-1}$ , using MEGAN. **Sindelarova2014** Estimate around  $594 \text{ Tg yr}^{-1}$  using MEGAN with MACC, showing isoprene as 69.2% of the total BVOC emissions, with monoterpenes at  $10.9 \text{ Tg yr}^{-1}$  (10.9%). They show  $41 \text{ Tg yr}^{-1}$  decrease in Australia when introducing soil moisture parameterisation.

When comparing the GEOS-Chem (which runs MEGAN) emissions to those calculated using our top-down inversion, we see a decrease over TODO: locations and seasons. TODO: table or figure showing summary of isoprene emissions changes over the whole of our time domain.

Satellite measured HCHO has been found to be biased low in several studies **Zhu2016**; **DeSmedt2015**; **Barkley2013** These papers use in-situ data to scale up the satellite HCHO columns for their areas of interest, however Australia lacks sufficient HCHO measurements to do this. In these papers bias is seen as high as 40%, which we use as our upper bound. Scaling up the satellite columns by 40% gives us an upper bound on the uncertainty due to satellite bias. The 'Scaled Satellite' column refers to the calculations when using the 40% scaled up OMI HCHO columns. This can be considered as a boundary on satellite based HCHO column uncertainty.

One set of data from the Daintree rainforest in Queensland exists (TODO: summary from P. Nelson). Although the data set lies outside our run times, as it was measured in TODO(runtime), we compare against the seasonal average of our GEOS-Chem output for the matching months (TODO: name the months). This is done for both GEOS-Chem output and our recalculated isoprene emissions. When compared against GEOS-Chem output we see TODO. When compared against recalculated emissions we see TODO.

TODO: Figure showing campaign data against model and recalculated emissions over region for averaged months and eventually different resolutions.

We examine the affect of decreased isoprene emissions on the correlation between modelled and satellite based HCHO columns. Figure TODO: shows the regressions between GEOS-Chem tropospheric column amounts of HCHO and satellite columns

for two runs of GEOS-Chem: a) using standard MEGAN emissions, b) using our updated emissions.

### 3.3.2 Emissions affect on GEOS-Chem

We interpolated or something (TODO) the emissions over Australia into the inventories used by GEOS-Chem which reduced the emissions by X% per year (over Australia). The resulting simulation output shows that HCHO was reduced by X%, although if we boost monoterpenes by X% where the isoprene emissions were lowered then

## 3.4 Uncertainty

There are several factors which need to be considered when looking at the uncertainty in emissions estimates. Things with their own inherent uncertainty include the modelled a-priori, modelled relationship between HCHO and isoprene, and satellite measurements. Important factors which need to be analysed for confidence in results include the steady state assumptions, filtering techniques for fire and human influences, and the regression model for determining the isoprene to HCHO yield.

Model uncertainty is difficult to accurately ascertain, generally an analysis of the model compared to in-situ measurements is performed, however there are few of these measurements over Australia.

### 3.4.1 Model Uncertainty

Uncertainty in modelled yield is estimated somehow (TODO:), and a upper and lower bound for the yield is determined using satellite scaling. Since OMI HCHO is scaled up by up to 40% in several papers, we consider HCHO scaled by 1 and 1.4 to be boundaries for modelled yield calculations. These prior works use flight campaigns and in-situ data to verify HCHO columns in various locations (TODO: redo cite list from lit review).

### 3.4.2 Satellite Uncertainty

Uncertainty in satellite measurements is generally provided along with the data, although uncertainty introduced through AMF calculation needs to be determined to give a representation of the confidence in vertical column amounts.

Provided with the OMI product is the measurement of uncertainty in each pixel, calculated by SAO from the backscattered solar radiation fit (Abad2015; Abad2016). BIRA use another method, and calculate the standard deviation of HCHO over the remote pacific ocean as the uncertainty (DeSmedt2012; DeSmedt2015). In the remote pacific, it can be assumed that HCHO variations are weak, with concentrations remaining steady in the short term ( $\sim 1$  month). This means the standard deviation over this region can be used as a proxy for determination of the instrument error.

There are three main sources of error in the resulting HCHO columns:

- a Fitting error from the OMI retrieval.

**b** Uncertainty in AMF calculations.

**c** Uncertainty of HCHO background.

a) is available in the OMI product and reduced through spatial and temporal averaging. Taking the eight day gridded average with horizontal resolution of 0.25 by 0.3125 degrees (latitude by longitude) typically reduces uncertainty by a factor of 1.5 to 4. Another method for examining uncertainty of OMI is to analyse the standard deviation of the HCHO columns over the remote pacific. If we assume there is no HCHO variation from background levels over any 8-day period, then this method infers variations in the measuring instrument, and can be used as a metric for uncertainty as done in **DeSmedt2012** TODO: uncertainty calculation on remote pacific OMI. **Millet2006**; **Palmer2006** both examine OMI HCHO columns over North America and determine overall uncertainty to be 40%, with most of this coming from cloud interference.

b) is determined through an analysis of GEOS-Chem output, validated against the total column of HCHO at Wollongong using FTIR measurements from the (TODO: Nicholas Jones roof HCHO citation here). **Palmer2006** calculate the error in AMF through combining estimates of error in the UV albedo database ( $\sim 8\%$ ), model error based on in-situ measurements, cloud error (20 – 30%) (**Martin2003**), and aerosol errors ( $< 20\%$ ), totalling AMF error of around  $\sim 30\%$ . It is worth noting here that independent error estimates are added in quadrature, which means total error equals the root of the sum of the independent errors each squared ( $e_{Total} = \sqrt{\sum_i e_i^2}$ ). TODO: Paul palmer calculation and combination for overall Satellite VC uncertainty per pixel and gridded. TODO: Millet2008?

c) is also determined through a study of GEOS-Chem output, in relation to in-situ measurements. TODO: calculate this uncertainty. Compare this error estimate with that of **Curci2010** where the error in b) and c) are respectively found to be 30% and 15% based on their analysis of CHIMERE. **Millet2008** also examine this uncertainty and determine an overall uncertainty ( $1\sigma$ ) of 25 – 27% in HCHO vertical columns with calculated AMFs where cloud fraction  $< 0.2$ .